

Predição de Níveis de Manutenção Preventiva em Motores de Aeronaves usando Técnicas de Aprendizado de Máquina

Leonardo de Souza Gomes¹

Ian Martins Arantes¹

¹Universidade Federal de Ouro Preto (UFOP)
Instituto de Ciências Exatas e Biológicas (ICEB)
Departamento de Computação (DECOM)
Ouro Preto, MG – Brazil

leonardo.sg@aluno.ufop.edu.br, ian.arantes@aluno.ufop.edu.br

Abstract. *This work presents an approach for predicting preventive maintenance levels in aircraft engines using machine learning techniques. Models such as Random Forest, Logistic Regression, Support Vector Machines (SVM), and Stochastic Gradient Descent (SGD) are employed to classify maintenance states into three categories: safe, preventive, and critical. A comparative analysis between the models is carried out based on classical evaluation metrics and the algorithms' execution time.*

Resumo. *Este trabalho apresenta uma abordagem para a predição dos níveis de manutenção preventiva em motores de aeronaves, utilizando técnicas de aprendizado de máquina. Empregamos modelos como Random Forest, Regressão Logística, Support Vector Machines (SVM) e Stochastic Gradient Descent (SGD) para classificar os estados de manutenção em três categorias: seguro, preventiva e crítica. A análise comparativa entre os modelos é realizada com base em métricas clássicas de avaliação e no tempo de execução dos algoritmos.*

1. Introdução

A manutenção preditiva é uma estratégia fundamental na indústria aeronáutica para garantir a segurança e reduzir custos operacionais, permitindo a realização de intervenções apenas quando necessárias. Este trabalho utiliza dados de sensores de motores de aeronaves para prever níveis de manutenção, classificados em três categorias: seguro, preventiva e crítica. Para tal, empregamos métodos de aprendizado de máquina supervisionados, que buscam identificar padrões nos dados que indicam o estado atual do motor.

2. Metodologia

Foram utilizados dados do conjunto *Predictive Maintenance Aircraft Engine*, disponibilizado no Kaggle, contendo séries temporais multivariadas de sensores para diferentes ciclos de operação dos motores. A variável alvo foi definida a partir do *Remaining Useful Life* (RUL), com classificação em três níveis:

- Crítica: $RUL \leq 30$ ciclos
- Preventiva: $30 < RUL \leq 80$ ciclos

- Seguro: RUL > 80 ciclos

Quatro modelos supervisionados foram treinados e avaliados: Random Forest, Regressão Logística, Support Vector Machines (SVM) e Stochastic Gradient Descent (SGD). Os dados foram pré-processados com seleção de features e normalização, quando necessário. Para cada modelo, foram medidos o tempo de treinamento e as métricas de desempenho classificatório.

A escolha dessas técnicas se justifica pelas características específicas de cada uma:

- **Random Forest:** algoritmo de ensemble baseado em múltiplas árvores de decisão, que oferece robustez e alta capacidade de generalização, além de lidar bem com dados complexos e variáveis correlacionadas.
- **Regressão Logística:** método probabilístico que permite uma interpretação clara dos coeficientes e serve como referência para modelos mais complexos.
- **Support Vector Machines (SVM):** eficaz para problemas com fronteiras de decisão não lineares.
- **Stochastic Gradient Descent (SGD):** técnica eficiente para otimização de modelos lineares em grandes conjuntos de dados, com treinamento rápido e escalável.

3. Métricas de Avaliação

Os modelos foram avaliados quanto às métricas: precisão, recall, F1-score e acurácia geral, além do tempo de execução.

Para medir o desempenho dos modelos, foram utilizadas as seguintes métricas:

- **Acurácia:** proporção de previsões corretas em relação ao total.
- **Precisão:** fração de exemplos positivos corretamente identificados dentre os positivos previstos.
- **Recall:** fração de exemplos positivos corretamente identificados dentre os exemplos positivos reais.
- **F1-Score:** média harmônica entre precisão e recall, equilibrando ambas.
- **Tempo de treinamento:** tempo gasto no treinamento de cada modelo, medido em segundos.

4. Resultados

Os resultados demonstraram que modelos como SVM e Random Forest obtiveram as melhores métricas, embora o Random Forest tenha apresentado maior tempo de treinamento. A Regressão Logística e SGD tiveram desempenho inferior, especialmente na classificação da classe preventiva, que é a mais desbalanceada. Limitações incluem o desbalanceamento dos dados, que pode impactar negativamente a performance das classes menos representadas.

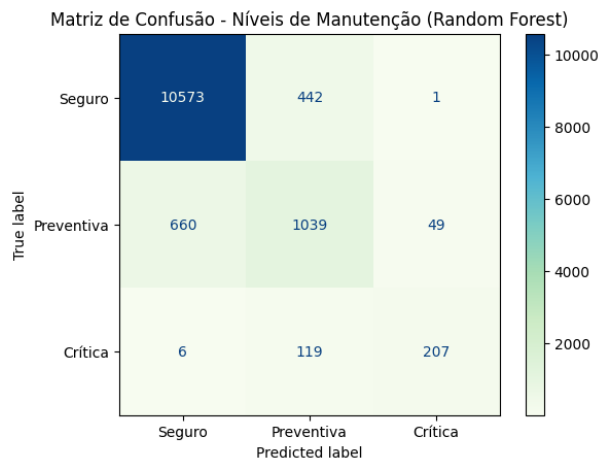


Figura 1. Matriz de Confusão - Random Forest

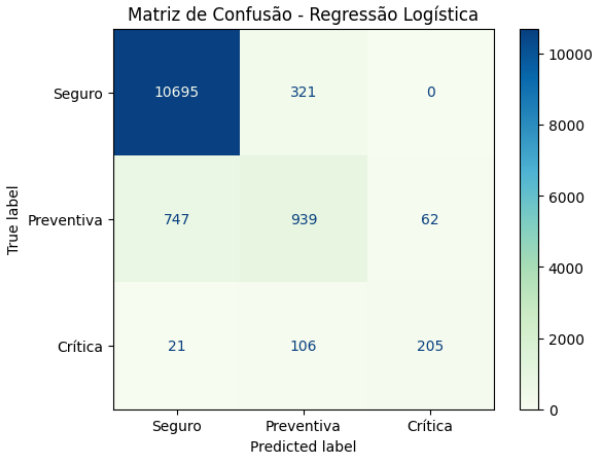


Figura 2. Matriz de Confusão - Regressão Logística

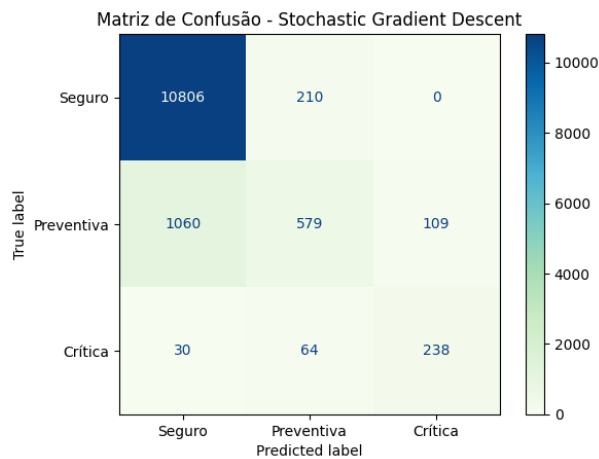


Figura 3. Matriz de Confusão - SGD

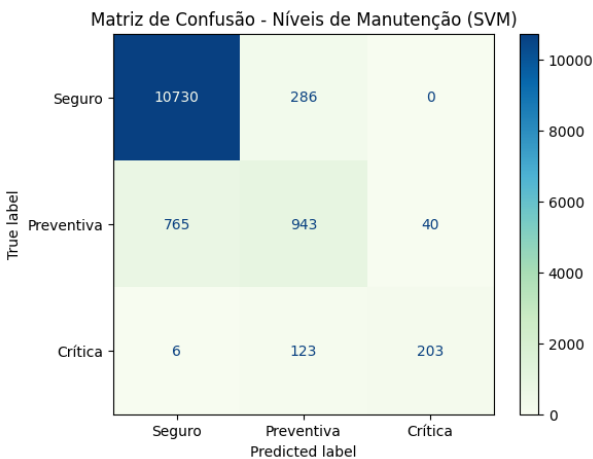


Figura 4. Matriz de Confusão - SVM

Modelo	Classe	Precisão	Recall	F1-Score	Acurácia Geral
Random Forest	Seguro	94.07%	95.98%	95.02%	90.25%
	Preventiva	64.94%	59.44%	62.07%	
	Crítica	80.54%	62.35%	70.29%	
Regressão Logística	Seguro	93.30%	97.09%	95.16%	90.40%
	Preventiva	68.74%	53.72%	60.31%	
	Crítica	76.78%	61.75%	68.45%	
SGD	Seguro	90.84%	98.09%	94.33%	88.75%
	Preventiva	67.88%	33.12%	44.52%	
	Crítica	68.59%	71.69%	70.10%	
SVM	Seguro	93.30%	97.40%	95.31%	90.68%
	Preventiva	69.75%	53.95%	60.84%	
	Crítica	83.54%	61.14%	70.61%	

Figura 5. Tabela de resultados dos modelos

Modelo	Precisão Média	Recall Médio	F1-Score Médio	Acurácia Geral	Tempo (s)
Random Forest	79.85%	72.59%	75.79%	90.25%	6.573
Regressão Logística	79.61%	70.85%	74.64%	90.40%	927
SGD	75.77%	67.63%	69.65%	88.75%	494
SVM	82.19%	70.83%	75.58%	90.68%	7.995

Figura 6. Tabela de resultados dos modelos

Em relação às *features*, foi observada uma melhora na avaliação dos resultados ao incluirmos a variável *cycle*, que indica em qual ciclo a turbina da aeronave se encontra. Embora ela não represente diretamente o RUL (*Remaining Useful Life*), possivelmente contribuiu para um melhor desempenho por fornecer uma noção mais clara de quão avançado está o uso do motor.

Outro comportamento observado foi no ajuste do parâmetro `n_estimators` do modelo *Random Forest*. Inicialmente, utilizamos 100 árvores e obtivemos resultados satisfatórios. Ao aumentar para 10.000 árvores, o ganho de acurácia foi muito pequeno e, ao elevar para 20.000 árvores, o incremento foi praticamente insignificante — cerca de 0,0001% na acurácia. Considerando que o tempo de treinamento com 100 árvores foi de 6,572 segundos, enquanto com 10.000 árvores chegou a quase dez minutos, optamos por manter o valor de `n_estimators` em 100, preservando o bom desempenho e reduzindo o custo computacional.

5. Conclusão

Os resultados indicam que todos os modelos alcançaram alta precisão e recall para a classe “Seguro”, evidenciando boa capacidade de identificar corretamente os casos sem necessidade imediata de manutenção. O destaque foi o modelo *SVM*, que obteve a maior acurácia geral (90,68%) e um equilíbrio notável entre precisão (93,30%) e recall (97,40%).

Na classe “Preventiva”, contudo, as métricas caem consideravelmente, especialmente no recall, o que revela dificuldade dos modelos em detectar essa categoria intermediária. O pior desempenho foi registrado pelo *SGD*, com recall de apenas 33,12%.

Para a classe “Crítica”, fundamental para prevenir falhas, os modelos *Random*

Forest e *SVM* apresentaram desempenho superior, com F1-score acima de 70%, sugerindo boa capacidade de identificar situações críticas, embora ainda haja espaço para avanços.

De forma geral, o *SVM* se sobressai por manter um equilíbrio consistente entre métricas e liderar em acurácia, enquanto *SGD* e *Regressão Logística* apresentam limitações, especialmente na detecção da classe “Preventiva”.

Como perspectivas futuras, recomenda-se investigar modelos voltados para séries temporais, como *ARIMA*, *LSTM* ou *Prophet*, que podem potencialmente aprimorar o desempenho preditivo e fornecer estimativas mais confiáveis e seguras para a manutenção aeronáutica.

6. Detalhes da Implementação

Os códigos foram desenvolvidos em Python, utilizando as bibliotecas `scikit-learn`, `pandas` e `matplotlib`. O repositório completo com o código fonte pode ser acessado em: <https://github.com/Leonardo2716ba/Predicao-de-Manutencao-em-Motores-de-Aeronaves>.

Os parâmetros principais usados foram:

- Random Forest: 100 árvores, `random_state=42`.
- Regressão Logística: `max_iter=100000`, `solver lbfgs`, `multinomial`.
- SVM: kernel RBF, `random_state=42`.
- SGD: `loss log_loss`, `max_iter=100000`, tolerância `1e-3`.

Referências

- Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer, New York.
- Murphy, K. P. (2012). *Machine Learning: A Probabilistic Perspective*. MIT Press, Cambridge, MA.
- Powers, D. M. W. (2011). Evaluation: From precision, recall and f-measure to roc, informedness, markedness and correlation. *Journal of Machine Learning Technologies*, 2(1):37–63.

[Murphy 2012] [Powers 2011] [Bishop 2006]